

MASTER'S THESIS

Beïnvloeden Seductive Details in Toetsitems de Interesse, Cognitieve Belasting, het Resultaat en de Benodigde Tijd?

Aalpoel, Anita

Award date:
2020

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 05. May. 2023

Open Universiteit
www.ou.nl



Beïnvloeden *Seductive Details* in Toetsitems de Interesse, Cognitieve Belasting, het Resultaat en de Benodigde Tijd?

Do Seductive Details in Testitems Influence Interest, Cognitive Load, Outcome and Time on Task?

F. J. Aalpoel

Master Onderwijswetenschappen
Open Universiteit

Datum: 21-02-2020

Naam begeleiders: Dr. K. J. H. Dirkx en Professor Dr. H. Jarodzka

Inhoudsopgave

Samenvatting.....	4
Summary	6
1. Inleiding	7
2. Theoretisch kader	8
2.1 Cognitieve theorieën over leren	8
2.2 Coherentieprincipe en leren	9
2.3 Typen <i>seductive details</i>	10
2.4 Coherentieprincipe bij toetsen.....	11
2.5 Vraagstellingen en hypothesen.....	13
3. Methode.....	14
3.1 Ontwerp.....	14
3.2 Participanten.....	15
3.3 Materialen	15
3.3.1 Rekentoets	15
3.3.2 Meten van interesse	18
3.3.3 Meten van de geleverde cognitieve belasting.....	18
3.4 Procedure.....	19
3.5 Data-analyse.....	20
4. Resultaten.....	21
4.1 Vooranalyse.....	21
4.2 Interesse.....	22
4.3 Cognitieve belasting.....	22
4.4 Toetsresultaat	22
4.5 Benodigde tijd	22
4.6 Post-hoc analyses	23
5. Conclusie en discussie.....	23
5.1 Conclusie.....	24
5.2 Discussie	24
5.2.1 Interesse.....	24
5.2.2 Cognitieve belasting.....	25
5.2.3 Toetsresultaat en tijd	26
5.3 Beperkingen van het onderzoek en aanbevelingen voor de praktijk.....	27

6. Referenties.....	28
Bijlagen	32
Bijlage 1. Verdeling van de toetsitems.....	32
Bijlage 2. Gemiddelden en standaard deviaties cognitieve belasting bij significant verschil op interesse.....	33
Bijlage 3. Gemiddelden en standaard deviaties toetsresultaat bij significant verschil op interesse..	34
Bijlage 4. Gemiddelden en standaard deviaties benodigde tijd bij significant verschil op interesse	35

Beïnvloeden *seductive details* in toetsitems de interesse, de cognitieve belasting, het resultaat en de benodigde tijd?

F.J. Aalpoel

Samenvatting

Afbeeldingen toevoegen aan leermaterialen blijkt effectief te zijn wanneer deze worden ontworpen volgens de principes van de *Cognitive Theory of Multimedia Learning* (CTML). Echter of deze principes ook gelden voor toetsen is nog weinig onderzocht. Afbeeldingen moeten effectief ingezet worden, zodat de cognitieve capaciteit optimaal benut kan worden. *Seductive details* bevatten irrelevante informatie en leiden af van de essentiële informatie, waardoor volgens het coherentieprincipe irrelevante cognitieve belasting wordt gecreëerd en leren gehinderd. Echter er zijn ook theorieën die stellen dat het toevoegen van *seductive details* zorgt voor een toename van de interesse, waardoor de lerende de aandacht vestigt op de inhoud en zich meer inspant.

In dit onderzoek staat centraal in hoeverre *seductive details* van toegevoegde waarde zijn in digitale toetsen, zodat er richtlijnen kunnen worden ontworpen voor het gebruik van deze details in relevante afbeeldingen. Doel van dit onderzoek is om zicht te krijgen op het effect van *seductive details* in relevante afbeeldingen bij een rekentoets in het primair onderwijs op de interesse, cognitieve belasting, het toetsresultaat en de benodigde tijd.

In dit *within-subject* experimenteel onderzoek hebben 67 leerlingen uit groep 7, tussen 9 en 10 jaar oud in het primair onderwijs, een authentieke rekentoets gemaakt, waarbij de helft van de items is aangepast volgens het coherentieprincipe en de andere helft niet. De toets is in een online omgeving gemaakt en na ieder toetsitem verschenen er twee vragen waarbij de leerling de interesse en cognitieve belasting per item beoordeelden.

De interesse is gemeten met een 4-punts beoordelingsschaal (Lindner, Ihme, Saß, & Köller, 2016) van 1 (“helemaal niet leuk”) tot 4 (“erg leuk”). Om de cognitieve belasting te meten is een 9-punts beoordelingsschaal (Paas, 1992) gebruikt waar de leerling de moeite die het kostte om het item te beantwoorden beoordeelden van 1 (‘zeer zeer weinig moeite’) tot 9 (‘zeer zeer veel moeite’). De toetsitems zijn geselecteerd uit de digitale rekentoets Groeimeter van Cito (2019).

De leerlingen blijken meer interesse te hebben in de toetsitems met *seductive details* dan in de toetsitems zonder deze *seductive details*. Er blijkt echter geen verschil te zijn in de geleverde cognitieve belasting, het resultaat en de benodigde tijd tussen de toetsitems zonder *seductive details* en de toetsitems met *seductive details*.

Op basis van de resultaten in huidig onderzoek lijkt het voor digitale toetsen dus niet zinvol om het coherentieprincipe toe te passen wanneer er relevante afbeeldingen worden gebruikt. Het tegenovergestelde kan zelfs geconcludeerd worden; het verdient aanbeveling om *seductive details* aan de relevante afbeeldingen in een toets toe te voegen, omdat dit zorgt voor een omgeving waarin

leerlingen meer geïnteresseerd zijn, zonder dat het de cognitieve belasting verhoogt of het resultaat en de benodigde tijd negatief beïnvloedt.

Trefwoorden: coherentieprincipe; *seductive details*; *Cognitive Theory of Multimedia Learning*; digitale toetsen; cognitieve belasting; primair onderwijs

Do seductive details in testitems influence interest, cognitive load, outcome and time on task?

F.J. Aalpoel

Summary

Adding images to learning resources appears affective if these images are designed in accordance with the principles of the Cognitive Theory of Multimedia Learning (CTML). Whether these principles also apply to tests has not yet been thoroughly researched. Images should be used effectively, so the cognitive capacity can be deployed optimally. Seductive details contain irrelevant information and detract from the essential information, which results in irrelevant cognitive load and obstruction of the learning process, according to the coherence principle. However, there are theories that suggest that adding seductive details may increase the interest, drawing the student's attention to the content and engaging them to make more effort.

The central question of this study is to what extent seductive details are of added value in digital tests, so guidelines can be designed for the use of these details in relevant images. The objective of this study is to gain insight into the effect of seductive details in relevant images on the interest, cognitive load, the outcome and time on task in an arithmetic test at primary education level.

In this within-subject experimental study, 67 students in group 7, ages between 9 and 10 in primary education, took an authentic arithmetic test, in which half the items had been changed in accordance with the coherence principle and the other half had not. The test was made in an online environment, and each test item was followed by two questions that allowed the student to evaluate their level of interest and cognitive load per item.

Interest levels were measured using a 4-point rating scale (Lindner, Ihme, Saß, & Köller, 2016) from 1 ("not fun at all") to 4 ("very fun"). The cognitive load has been measured using a 9-point rating scale (Paas, 1992) in which the student rated the effort it took them to answer the item from 1 ('very, very little effort') to 9 ('very, very much effort'). The test items were selected from the Cito Groeimeter (2019) digital arithmetic test.

The students appeared more interested in the test items with seductive details than the test items without these seductive details. However, there appears to be no difference in the resulting cognitive load, the outcome and the time on task between the test items without seductive details and the test items with seductive details.

Based on the results of this study, it does not seem useful to apply the coherence principle in digital tests if relevant images are used. In fact, the opposite can be concluded; it is advisable to add seductive details to the relevant images in a test, as this results in an environment that is more interesting to students, without increasing cognitive load or negatively influencing the outcome or time on task.

Keywords: coherence principle; seductive details; Cognitive Theory of Multimedia Learning; digital tests; cognitive load; primary education

1. Inleiding

De huidige generatie digitale toetsen zijn vaak gebaseerd op traditionele ontwerpprincipes voor papieren versies. Terwijl papieren toetsen slechts een beperkte mogelijkheid hebben voor het toevoegen van media, zoals afbeeldingen, bieden digitaal toetsen veel mogelijkheden voor het toevoegen van multimedia, zoals (interactieve) afbeeldingen, video en audio. Het gebruik van multimedia in digitale toetsen leidt echter tot nieuwe ontwerp vragen waarvoor tot op heden nog geen empirisch gevalideerde richtlijnen beschikbaar zijn (Kirschner, Park, Malone, & Jarodzka, 2016).

Uit onderzoek naar het effect van afbeeldingen bij leren blijkt dat leren met woorden én afbeeldingen effectiever is dan leren met alleen woorden, in elk geval als er rekening wordt gehouden met de principes van *Cognitive Theory of Multimedia Learning* (CTML; Mayer, 2014). De CTML zou volgens Beddow (2018) en Kirschner et al. (2016) ook relevant kunnen zijn voor het ontwerpen van digitale toetsen waarin multimediaal materiaal wordt gebruikt.

Eén van de CTML principes is het coherentieprincipe. Het uitgangspunt van dit principe is, dat leermaterialen optimaler te verwerken zijn en het leren effectiever maken, zonder toevoeging van irrelevante informatie (Valcke, 2010). Wat hier gezien wordt als irrelevante informatie, de zogenaamde *seductive details*, is informatie die geen relatie heeft met het leerdoel en zorgt voor een cognitieve overbelasting die het leren hindert (Mayer, 2014). Het toepassen van het coherentieprincipe bij het ontwikkelen van leermateriaal zorgt ervoor dat het leren wordt bevorderd doordat alle cognitieve capaciteit ingezet kan worden voor essentiële kennisverwerking.

Hoewel het coherentieprincipe in leermateriaal veelvuldig is onderzocht (Butcher, 2006; Mayer & Fiorella, 2014; Mayer, Griffith, Jurkowitz, & Rothman, 2008; Park, Flowerday, & Brünken, 2015; Sung & Mayer, 2012; Wang & Adesope, 2014), is er over het elimineren van *seductive details* bij toetsen nog onvoldoende bekend (Lindner, Eitel, Strobel, & Köller, 2017a). Wel wordt er in de praktijk juist veel gebruik gemaakt van afbeeldingen met veel kleur, het toevoegen van authentieke elementen, zoals personen, etc. met het idee dat dit zorgt voor een grotere interesse tijdens het maken van de opgaven (zeker bij formatieve toetsen). Vanuit het coherentieprincipe is het maar de vraag of het toevoegen van dit soort informatie de toetsitems niet onnodig moeilijk maakt.

Dit onderzoek richtte zich dan ook op de vraag of het toevoegen van *seductive details* bij het ontwerpen van digitale toetsen in het primair onderwijs een effect heeft op de interesse, de cognitieve belasting, het toetsresultaat en de benodigde tijd om het item te beantwoorden. Er is hierbij gekeken naar het gebruik van afbeeldingen, omdat deze het meest gebruikt worden in digitale toetsen en vaak veel irrelevante informatie bevatten. Andere vormen van multimedia, zoals video en audio, worden in het primair onderwijs nog maar op beperkte schaal toegepast in digitale toetsen. De resultaten uit huidige onderzoek kunnen helpen om de vraag naar de ontwikkeling van goede richtlijnen voor het gebruik van afbeeldingen in digitale toetsen te beantwoorden.

2. Theoretisch kader

Door het toevoegen van afbeeldingen aan leermaterialen willen ontwerpers een leeromgeving creëren die prettig oogt voor de leerlingen en hun aanspreekt. Verder heeft technologie ervoor gezorgd dat afbeeldingen een steeds grotere rol spelen bij toetsen op de computer, omdat er vele technische mogelijkheden zijn om deze toe te voegen aan toetsmateriaal (Lindner, Lüdtke, Grund & Köller, 2017b). Er wordt daarbij door de ontwerpers echter onvoldoende gekeken naar de werking van het menselijke informatie verwerkingssysteem (Mayer, 2014). Beddow (2018) en Kirschner et al. (2016) geven dan ook aan dat bij het gebruiken van afbeeldingen in toetsen, de theorieën die hierover bekend zijn, aan de basis zouden moeten liggen van het ontwerp om de validiteit te waarborgen.

2.1 Cognitieve theorieën over leren

Er zijn twee prominente theorieën die vaak gebruikt worden als het gaat om het ontwerpen van leermateriaal met afbeeldingen. Dit zijn de *Cognitive Theory of Multimedia Learning* (CTML) van Mayer (2014) en de *Cognitive Load Theory* (CLT) van Sweller, Van Merriënboer, en Paas (1998). De onderliggende ideeën van beide theorieën lijken erg veel op elkaar en worden daarom hieronder toegelicht.

De CTML en CLT zijn gebaseerd op drie cognitieve principes van leren (Mayer, 2014). Als eerste de *dual-channel assumption* (Paivio, 1990), waarbij het uitgangspunt is dat mensen twee kanalen hebben om informatie te verwerken: het visuele verwerkingskanaal en het auditieve/verbale verwerkingskanaal. Bij het leren met multimedia worden beide kanalen geactiveerd. Geschreven of gesproken woorden activeren het auditieve kanaal en afbeeldingen het visuele kanaal. Het tweede principe waar de theorieën op gebaseerd zijn is de gelimiteerde capaciteit veronderstelling, waarbij het uitgangspunt is dat ieder verwerkingskanaal een beperkte capaciteit heeft (Baddeley, 1999). Wanneer beide kanalen ingezet worden, kan er dus meer informatie gelijktijdig verwerkt worden. Het derde principe gaat er van uit dat leren alleen zinvol is wanneer er actieve verwerking van de gepresenteerde informatie plaatsvindt. Er zijn drie cognitieve processen nodig om actieve verwerking te laten plaatsvinden. Bij 'selecteren' vindt er een transfer plaats van relevante informatie naar het werkgeheugen, bij 'organiseren' worden er relaties gelegd en ontstaat er een samenhangende cognitieve structuur in het werkgeheugen en bij 'integratie' wordt deze structuur gekoppeld aan relevante voorkennis die is geactiveerd vanuit het lange termijngeheugen.

Verder gaan beide theorieën ervan uit dat de cognitieve capaciteit beïnvloed kan worden door twee soorten belasting (Sweller, Ayres, & Kalyuga, 2011). De eerste is de *intrinsic cognitive load*, waarbij het gaat om de complexiteit van de taak die bepaald wordt door het aantal elementen dat geleerd moeten worden en de interactie, ofwel de samenhang, tussen deze elementen. Tevens gaat het om de aanwezige voorkennis bij de lerende, omdat een expert een andere complexiteit ervaart dan een beginner. De *intrinsic cognitive load* hangt dus af van een combinatie van zowel de aard van de informatie als de kennis van de persoon die de informatie verwerkt (Sweller, Van Merriënboer, &

Paas, 2019). De *intrinsic cognitive load* kan worden verlaagd door de taak te veranderen of de kennis van de lerende te vergroten. De andere soort belasting is de *extraneous cognitive load*. Het betreft de irrelevante belasting veroorzaakt door het ontwerp van het leermateriaal en belemmert het leren (Sweller et al., 2011). Verhoging van de *extraneous cognitive load* gebeurt wanneer het materiaal niet op de juiste wijze is gepresenteerd, de instructie onduidelijk is of doordat het irrelevante informatie bevat. Het hangt samen met de *intrinsic cognitive load*, omdat effectieve instructie de interactie tussen de eerdere genoemde elementen kan verminderen (Sweller, 2010) en hoe meer middelen moeten worden besteed aan het omgaan met de *extraneous cognitive load* hoe minder er beschikbaar zal zijn voor het omgaan met de *intrinsic cognitive load*, met als gevolg dat er dus minder wordt geleerd (Sweller et al., 2019). Het is dus belangrijk om deze *extraneous cognitive load* te verlagen door het materiaal zo te ontwerpen dat het optimaal aan de uitgangspunten van de CTML is aangepast. Het doel van instructieontwerp volgens de CTML principes is in lijn met de CLT. Het gaat om het optimaliseren van *intrinsic cognitive load* en het verminderen van *extraneous cognitive load* (Mayer, 2014) door het effectief gebruiken van afbeeldingen (of andersoortige multimedia). Het hieronder toegelichte coherentieprincipe kan hieraan bijdragen.

2.2 Coherentieprincipe en leren

Hoewel één van de basisprincipes (i.e. het multimediateprincipe) van CTML stelt dat het toevoegen van afbeeldingen aan het leermateriaal bijdraagt aan het leren, garandeert het eenvoudigweg toevoegen van een afbeelding niet dat er ook echt beter wordt geleerd (Mayer, 2014). De afbeelding moet voldoen aan het zogeheten coherentieprincipe (Mayer & Fiorella, 2014). Volgens dit principe moet alle irrelevante informatie – zoals afbeeldingen, woorden en geluiden die niet bijdragen aan het begrijpen van leerstof – vermeden worden. Het coherentieprincipe kan dus helpen om vanuit een meer genuanceerde benadering naar het gebruik van afbeeldingen in leermateriaal te kijken. Dit is noodzakelijk om te bepalen hoe afbeeldingen effectief ingezet kunnen worden, zodat het een gunstig effect heeft op de leerresultaten (Butcher, 2014).

Relevante afbeeldingen verlagen de *extraneous cognitive load*. Het zijn afbeeldingen die noodzakelijk zijn om de taak te kunnen volbrengen. Ze verbeteren de mentale representatie die de lerende construeert (Lindner et al., 2017b). Schnotz (2014) bevestigt dat afbeeldingen het leren bevorderen als ze gerelateerd zijn aan de tekst die wordt gepresenteerd. *Seductive details* daarentegen bevatten irrelevante informatie die toegevoegd zijn aan deze relevante afbeeldingen. Ze zijn erg aandachttrekkend en kunnen als zeer interessant ervaren worden door de lerende, ze zijn echter niet direct gerelateerd aan het instructiedoel. Doordat deze *seductive details* afleiden van de essentiële informatie, wordt volgens het coherentieprincipe de *extraneous cognitive load* onnodig verhoogd en het leren gehinderd (Sung & Mayer, 2012).

Butcher (2006) heeft in zijn onderzoek naar lessen over het menselijke hart bevestigd dat simpele, relevante afbeeldingen zonder *seductive details* zorgen voor een beter begrip van de leerstof. Mayer en

Fiorella (2014) tonen dit ook aan in vele onderzoeken naar begrip over bliksem en laten zien dat *seductive details* in afbeeldingen het leren belemmeren. Mayer et al. (2008) hebben onderzoek gedaan naar de uitleg over een verkoudheidsvirus, die het menselijk lichaam infecteert, en over de werking van onze spijsvertering. In beide experimenten nam het begrip van de lerenden af naarmate zij de details in de relevante afbeeldingen interessanter vonden. Onderzoek beperkte zich echter veelal tot de natuurwetenschappen. Wang en Adesope (2014) hebben daarom onderzoek gedaan naar het coherentieprincipe in de sociale wetenschappen. Ook in dat onderzoek blijkt dat het elimineren van *seductive details* effectief is. In hun onderzoek, waarbij het doel was meer te leren over de economische situatie in een willekeurige staat van Amerika, presteerde de groep zonder *seductive details* beter dan de groep die informatie had ontvangen met *seductive details*.

Toch worden *seductive details* veelvuldig toegevoegd aan leer- en toetsmateriaal met als doel het materiaal interessanter en aantrekkelijker te maken (Wang & Adesope, 2014). Onderzoek laat namelijk ook zien dat *seductive details* de interesse van de lerenden wekt (Towler et al., 2008). Izard en Ackerman (2000) onderstrepen het belang van interesse, omdat daarmee motivatie voor leren en de persoonlijke betrokkenheid zouden worden vergroot. De *Cognitive Affective Theory of Learning with Media* (CATLM; Moreno, 2005) veronderstelt ook dat lerenden interesse moeten hebben in de leermaterialen om volledig gebruik te maken van hun cognitieve middelen. Interessant materiaal stimuleert de lerende, waardoor de aandacht wordt gevestigd op de inhoud en de inspanning wordt verhoogd. Ook in het onderzoek van Park et al. (2015), waarbij instructie werd gegeven op cellulaire moleculen in een online omgeving, hebben *seductive details* geleid tot gerichte aandacht en een affectieve reactie die leidde tot een leerproces van hogere kwaliteit. Deze bevindingen staan echter in contrast met het coherentieprincipe, waarbij wordt gesteld dat *seductive details* afleiden van de inhoud en het leren hinderen.

2.3 Typen *seductive details*

Er worden in de literatuur verschillende kenmerken of aspecten aangemerkt als zijnde *seductive details*. De meeste literatuur richt zich daarbij op kleur en menselijke kenmerken. Door relevante afbeeldingen te voorzien van kleuren en menselijke kenmerken (bijvoorbeeld gezichten met expressieve ogen) kan de interesse voor het leermateriaal worden verhoogd. Dit blijkt uit onderzoek van Mayer en Estrella (2014) naar begrip over het ontstaan van virussen. De groep die afbeeldingen ontving met toegevoegde menselijke kenmerken en kleur hadden betere leerresultaten, doordat de interesse van de lerende toenam. Echter uit meer recent onderzoek door Stárková, Lukavsky, Javora, en Brom (2019) naar de effecten van het toevoegen van menselijke kenmerken en kleur aan leermateriaal bleek dat dit geen effect heeft op leerresultaten. Clark en Mayer (2016) hebben zelfs geconcludeerd in eerder onderzoek over begripsvorming over het ontstaan van bliksem, dat de groep die uitleg kreeg zonder toevoeging van kleuren beter presteerde dan de groep met gekleurde afbeeldingen. Dit laatste onderzoek is in lijn met het coherentieprincipe, omdat de menselijke

kenmerken en kleuren *seductive details* zijn die zorgen voor een verhoging van de *extraneous cognitive load*, waardoor te weinig cognitieve belasting beschikbaar is in het werkgeheugen om de essentiële informatie te verwerken.

Harp en Mayer (1998) geven drie verklaringen waarom *seductive details* het leren (kunnen) hinderen. De eerste noemen ze de afleidingshypothese; de *seductive details* trekken de aandacht weg van de belangrijke informatie. De tweede verklaring wordt de verstoringshypothese genoemd; *seductive details* onderbreken de transitie van het ene hoofdidee naar het andere en er wordt te veel tijd besteed aan irrelevante informatie, waardoor diepere verwerking niet ontstaat. De laatste verklaring is de zogenaamde omleidingshypothese, waarbij de lerende een mentale representatie opbouwt rondom de *seductive details* in plaats van om de belangrijke informatie. Dit ontstaat doordat de *seductive details* in de afbeelding zorgen voor inadequate actieve verwerking. Tijdens het proces van ‘selecteren’ wordt irrelevante informatie geselecteerd, hierdoor worden er bij de fase van het ‘organiseren’ verkeerde relaties gelegd, wat een negatieve invloed heeft op het ontstaan van samenhangende cognitieve structuren en vervolgens staat het de ‘integratie’ in de weg, omdat verkeerde voorkennis uit het lange termijngeheugen is geactiveerd door de irrelevante informatie. Rey (2012) bevestigt op basis van een meta-analyse deze omleidingshypothese.

Seductive details kunnen dus leiden tot een hogere cognitieve belasting, omdat er veel irrelevante informatie verwerkt moet worden met als resultaat dat er minder cognitieve capaciteit beschikbaar is voor het leren (Mayer et al., 2008; Sung & Mayer, 2012; Wang & Adesope, 2014). Anderzijds laat de literatuur zien dat het toevoegen van *seductive details* kan leiden tot een grotere interesse (Izard & Ackermann, 2000; Moreno, 2005; Towler et al., 2008) en betere prestaties (Mayer & Estrella, 2014; Park et al., 2015). Doordat het toevoegen van *seductive details* aan relevante afbeeldingen niet bevorderlijk kan zijn of zelfs negatieve effecten kan hebben, omdat ze de constructie van mentale representaties verstoren (De Westelinck, Valcke, De Craene, & Kirschner, 2005), kan het volgens het coherentieprincipe zinvol zijn om het toevoegen van *seductive details* te vermijden (Mayer, 2014). Beide perspectieven in ogenschouw nemend dient er dus een weloverwogen besluit te worden genomen welke visualisatie het meest geschikt lijkt voor verschillende typen materialen en doelen (Schnotz & Bannert, 2003).

2.4 Coherentieprincipe bij toetsen

Het effect van de CTML en het bijbehorende coherentieprincipe op leren zijn dus veelvuldig onderzocht, maar hoe zit het met het coherentieprincipe bij toetsen? De doelstelling van toetsen is tenslotte anders, dus hoewel gesuggereerd wordt dat de CTML principes toepasbaar zijn op toetsen, is dat nog maar de vraag (Jarodzka, Janssen, Kirschner, & Erkens, 2015; Kirschner et al., 2016). Het gaat bij toetsen namelijk om het beoordelen of evalueren van de mate waarin een lerende zijn of haar kennis en vaardigheden heeft verworven, behouden en/of kan toepassen. Terwijl de doelstelling van leren is om deze kennis te verwerven, te behouden en toe te leren passen (Beddow, 2018).

Uit eerder onderzoek naar digitale toetsen vanuit multimedia perspectief is bekend dat leerlingen betere toetsresultaten behalen wanneer een toets wordt gemaakt met woorden én afbeeldingen dan met alleen woorden (multimediaprincipe; Lindner et al., 2017a; Lindner et al., 2017b). Deze onderzoeken zijn echter vooral gericht op het toevoegen van relevante afbeeldingen aan toetsen en laten zien dat de moeilijkheid van een toetsitem afneemt en toetsen succesvoller worden gemaakt wanneer relevante afbeeldingen worden toegevoegd. Uit onderzoek van Lindner, Ihme, Saß, en Köller (2016) blijkt verder dat relevante afbeeldingen zorgden voor een toename in het plezier bij de lerende en daardoor volgens hen vaker moet worden overwogen om afbeeldingen te gebruiken.

Er is daarbij niet gekeken naar de hoeveelheid relevante informatie in de afbeeldingen en het effect van *seductive details* in deze relevante afbeeldingen op de toetsprestaties (Lindner et al., 2017a), terwijl (in elk geval in de toetsen van Cito) vaak gebruikt gemaakt wordt van aantrekkelijke afbeeldingen met veel *seductive details* zoals kleuren, menselijke kenmerken etc. Vanuit onderzoek van Ögren, Nyström, en Jarodzka (2017) is bekend dat afbeeldingen in toetsen over het algemeen heel erg de aandacht trekken en afleiden van de gegeven schriftelijke informatie in de vraag. Afbeeldingen kunnen zelfs misleiden als de gegeven informatie in de afbeelding door de lerende zonder meer wordt geloofd. Het zou dus logisch zijn als de hoeveelheid *seductive details* tot een minimum beperkt wordt, zodat de lerende zich kan focussen op het begrijpen en beantwoorden van de vraag. Onderzoek van Manastirean-Zijlstra (2017) bevestigt dat het elimineren van irrelevante visuele informatie in een rekentoets zorgt voor een verhoging in de prestatie doordat de lerenden deze afbeeldingen niet hoeven te verwerken en er daarom geen cognitieve capaciteit verloren gaat (of de verhoging van de prestatie echter terug te voeren is naar de toepassing van het coherentieprincipe is onduidelijk, aangezien meerdere principes van de CTML werden gecombineerd).

Deze onderzoeken richtten zich echter allemaal op toetsen waarbij de afbeelding in zijn geheel overbodig is. Onderzoek naar het effect van het elimineren van *seductive details* in relevante, niet overbodige afbeeldingen op de interesse, cognitieve belasting, het toetsresultaat en de benodigde tijd is tot dusver dus nog beperkt, zij het niet volledig afwezig. Zoals we hiervoor gezien hebben kan het toevoegen van *seductive details* leiden tot meer interesse voor de materialen, waardoor het een positief effect heeft, maar tegelijkertijd kunnen de *seductive details* in de afbeeldingen juist ook zorgen voor een onnodige verhoging van de *extraneous cognitive load* en dus voor een negatief effect. Beddow (2018) heeft in literatuuronderzoek bekeken in hoeverre het coherentieprincipe bij relevante afbeeldingen gebruikt kan worden voor het ontwerpen van toetsen. Als reden voor het toevoegen van *seductive details* aan relevante afbeeldingen geeft Beddow (2018) aan dat dit kan zorgen voor een representatie die meer op de echte wereld lijkt, omdat de context van een probleem uit de echte wereld wordt gegeven. De toevoeging van zulke *seductive details*, zodat de opgave meer lijkt op een probleem uit de echte wereld, zal volgens hem niet leiden tot een incorrect antwoord, maar het kan ervoor zorgen dat degene die de toets maakt onnodig veel tijd en/of cognitieve middelen in moet zetten die wellicht nodig zijn voor het oplossen van het item. Dit is in lijn met de verstoringshypothese

bij leren van Harp en Mayer (1998) die tevens stellen dat er te veel tijd wordt besteed aan de irrelevante informatie. Beddow (2018) geeft als aanbeveling toetsitems te vereenvoudigen door *seductive details* in de relevante afbeelding te verwijderen. De *extraneous cognitive load*, die door deze *seductive details* wordt gecreëerd, moet worden geëlimineerd. Door deze *extraneous cognitive load* te elimineren zijn verschillen in toetsresultaat niet het gevolg van verschillen in werkgeheugencapaciteit van de getoetste persoon en wordt met de toets alleen het leren van de persoon geëvalueerd of beoordeeld. Beddow (2018) heeft echter niet onderzocht of de aanbevelingen voor toepassing van het coherentieprincipe bij toetsen daadwerkelijk bijdragen aan een verbetering van de toetsen. Hij doet slechts suggesties voor het ontwerp van een toets op basis van literatuuronderzoek over het coherentieprincipe. Meer onderzoek is dus nodig om vast te kunnen stellen of het toepassen van het coherentieprincipe, waarbij *seductive details* in relevante afbeeldingen zijn geëlimineerd, daadwerkelijk een positief effect heeft op de geleverde cognitieve belasting en de toetsresultaten en wat het effect hiervan is op de interesse van de lerende en de benodigde tijd om het item te beantwoorden.

2.5 Vraagstellingen en hypothesen

Gezien bovenstaande contrasterende bevindingen en het feit dat er in veel toetsitems (van Cito) wel *seductive details* terug te vinden zijn, hebben we in samenwerking met Cito huidig onderzoek uitgevoerd. Doel van dit onderzoek was om zicht te krijgen op het effect van *seductive details* bij een rekentoets in het primair onderwijs op de interesse, cognitieve belasting, het toetsresultaat en de benodigde tijd, zodat de resultaten kunnen bijdragen aan het ontwerpen van richtlijnen voor het gebruik van *seductive details* in relevante afbeeldingen van digitale toetsen. De volgende vraag stond hierom centraal:

Wat is het effect van seductive details bij relevante afbeeldingen op de interesse, de cognitieve belasting, het toetsresultaat en de benodigde tijd bij een digitale rekentoets in het primair onderwijs?

Seductive details worden vaak toegevoegd om te zorgen voor een betere representatie van de echte wereld (Beddow, 2018), waardoor de interesse voor het toetsitem toeneemt (Wang & Adesope, 2014). De eerste hypothese luidde daarom als volgt:

H₁: Er is een verschil in interesse die de lerende heeft voor de toetsitems met relevante afbeeldingen zonder *seductive details* en de items met relevante afbeeldingen met *seductive details*. De lerenden zijn meer geïnteresseerd in de items met *seductive details*, dan in de items zonder *seductive details*.

Op basis van onderzoek naar het coherentieprincipe bij leren lijkt het echter zinvol om *seductive details* te elimineren, omdat het gebruik ervan irrelevante cognitieve belasting creëert. Er gaat cognitieve belasting verloren, waardoor er minder beschikbaar is om te leren (Sung & Mayer, 2012; Wang & Adesope, 2014). Beddow (2018) en Kirschner et al (2016) geven als aanbeveling het

coherentieprincipe ook toe te passen bij toetsen, zodat de beperkte cognitieve capaciteit volledig ingezet kunnen worden om het toetsitem juist te beantwoorden. Dit leidde tot de volgende hypothese: H₂: Er is een verschil in cognitieve belasting tussen de toetsitems met relevante afbeeldingen zonder *seductive details* en de toetsitems met relevante afbeeldingen met *seductive details*. De toetsitems zonder *seductive details* zorgen voor een lagere zelfgerapporteerde cognitieve belasting dan toetsitems met *seductive details*.

Doordat de *seductive details* erg aandachttrekkend zijn en het leren kunnen hinderen (Sung & Mayer, 2012) is de verwachting dat er op deze toetsitems minder wordt gepresteerd. Bewijs vanuit de CTL (Sweller et al., 1998) en CTML (Mayer, 2014), dat *seductive details* door de verhoging van de cognitieve belasting zorgen voor een lager toetsresultaat, is overtuigender dan het bewijs om *seductive details* te behouden vanuit interesse perspectief. Er is tenslotte meer onderzoek gedaan naar de CTL en CTML en deze zijn qua opzet overtuigender dan het beperkte en veelal theoretische onderzoek naar het effect van interesse op de resultaten. Dit leidde tot de derde hypothese:

H₃: Er is een verschil in toetsresultaat van de toetsitems met relevante afbeeldingen zonder *seductive details* en de items met relevante afbeeldingen met *seductive details*. Toetsitems zonder *seductive details* worden beter gemaakt dan toetsitems met *seductive details*.

Beddow (2018) stelt dat toetsitems met *seductive details* vooral onnodig veel tijd van de lerende kosten om de items te beantwoorden en dat deze details daarom geëlimineerd moeten worden. Dit heeft geleid tot de laatste hypothese:

H₄: Er is een verschil in de benodigde tijd om de toetsitems met relevante afbeeldingen zonder *seductive details* te beantwoorden en de items met relevante afbeeldingen met *seductive details*. Toetsitems met *seductive details* kosten meer tijd om te beantwoorden dan de toetsitems zonder *seductive details*.

3. Methode

Dit onderzoek was onderdeel van de thesiskring van de faculteit Onderwijswetenschappen met als thema *Instructional design for online learning*. Binnen deze onderzoekslijn ligt de focus op het ontwikkelen van richtlijnen voor digitale toetsen. Er wordt onderzocht of de principes van de CTML (Mayer, 2014) ook gelden voor digitaal toetsen. De voorbereidingen voor huidig onderzoek werden samen gedaan met een andere student, die het modaliteitsprincipe onderzocht. De data werd afzonderlijk van elkaar verzameld. Andere principes uit de CTML werden en worden door andere studenten binnen reeds afgesloten en lopende thesiskringen onderzocht (i.e., Manastirean-Zijlstra, 2017).

3.1 Ontwerp

In deze masterthesis werd een experimenteel *within-subject* onderzoek uitgevoerd (Creswell, 2014) met type afbeelding als onafhankelijke *within-subject* variabele en interesse, cognitieve belasting,

toetsresultaat en benodigde tijd als afhankelijke variabelen om binnen-persoonsvariabelen (Kalyuga, Ayres, Chandler, & Sweller, 2003) gelijk te verdelen over de condities. Het zou voor de deelnemers te belastend geweest kunnen zijn om naast het beantwoorden van het item, het beoordelen van de interesse en de cognitieve belasting ook nog de voorkennis en het ruimtelijk inzicht te meten. Alle deelnemers beantwoordden dus zowel items zonder *seductive details* als items met *seductive details*.

3.2 Participanten

Deelnemers aan het onderzoek waren leerlingen uit drie groepen 7 van een basisschool in het midden van Nederland. De leeftijd van de leerlingen was tussen de 9 en 10 jaar ($M = 9.94$, $SD = 0.24$). De steekproef werd niet random getrokken, want de directeur van de school was een contact uit het netwerk van de onderzoeker. Er werd gestreefd naar een minimale deelname van 44 leerlingen. Dit werd gebaseerd op een powerberekening, waarbij uit is gegaan van $\alpha = .05$ en een effectgrootte van $r = .56$ (Manastirean-Zijlstra, 2017). Uiteindelijk namen 67 participanten deel aan het onderzoek. Oorspronkelijk waren 70 participanten benaderd om deel te nemen, maar voor drie leerlingen meldden de betreffende ouders dat ze geen toestemming gaven voor deelname. De participanten werden a-select systematisch verdeeld over de twee toetsversies door de leerlingcodes van de toetsversies zo te verdelen dat steeds het uitdelen van een leerlingcode van versie 1 werd afgewisseld met het uitdelen van een leerlingcode van versie 2. Hierdoor waren de toetsversies 1 en 2 vrijwel gelijk verdeeld over de groepen. In tabel 1 is het aantal participanten, de gemiddelde leeftijd en de standaarddeviatie weergegeven.

Tabel 1
Aantal Participanten en Gemiddelde Leeftijd per Toetsversie

	<i>n</i>	<i>M</i>	<i>SD</i>
Totaal	67	9.94	0.24
Toetsversie 1	35	9.97	0.17
Toetsversie 2	32	9.91	0.30

3.3 Materialen

Voor het huidige onderzoek werd gebruikt gemaakt van opgaven uit een bestaande rekentoets van Cito (2019), de 4-punts beoordelingsschaal voor interesse van Lindner et al. (2016) en de 9-punts beoordelingsschaal voor cognitieve belasting van Paas (1992).

3.3.1 Rekentoets

Bij de leerlingen werd een rekentoets afgenomen, waarbij de relevante afbeeldingen werden gemanipuleerd door deze aan te passen aan het coherentieprincipe (Mayer, 2014). De helft van de rekenopgaven bevatte relevante afbeeldingen zonder *seductive details*, de andere helft bevatte

relevante afbeeldingen met *seductive details*. Beiden werden door alle participanten gemaakt. Er werden twee versies van de toets ontwikkeld, zodat voorkomen kon worden dat de onderzoeksresultaten toegeschreven konden worden aan de eventuele toevallige complexiteit van het toetsitem, ongeacht of deze aangepast was volgens het coherentieprincipe.

De toetsitems die voor dit onderzoek werden gebruikt, zijn een selectie van items uit het online platform Groeimeter van Cito (2019) van jaargroep 6. De items van groep 6 werden getoetst in groep 7, doordat het onderzoek werd afgenomen na slechts acht weken onderwijs in het nieuwe schooljaar. Leerlingen uit groep 6 zouden nog onvoldoende instructie hebben gehad over de toetsdoelen en onvoldoende kennis om de vragen te kunnen beantwoorden. Er zou dan dus niet geëvalueerd of beoordeeld zijn of de verworven kennis en vaardigheden was behouden en werd toegepast, aangezien ze deze nog niet hadden verworven.

Groeimeter is een digitale rekentoets voor formatieve assessment ontwikkeld door Cito (2019) voor het primair onderwijs. De Citotoetsen zijn gecertificeerd door de Commissie Testaangelegenheden Nederland (COTAN; Sanders et al., 2017). Groeimeter (Cito, 2019) bevat meer dan 500 items voor de verschillende leerdoelen voor rekenen in groep 6. Uit deze 500 werden in drie rondes door de drie betrokken onderzoekers 26 items geselecteerd. Er werd gekozen voor 26 items, omdat uit andere Cito rekentoetsen blijkt dat dit een reëel aantal is voor een aaneengesloten afname. De selectie werd gedaan op basis van de volgende criteria: 1) De items bevatten een relevante afbeelding die nodig is om de vraag te kunnen beantwoorden. Dit is niet bij alle items in de Groeimeter het geval. Er zijn ook items met afbeeldingen die beantwoord kunnen worden zonder de afbeelding te gebruiken. Aangezien dit onderzoek ging over het gebruik van relevante afbeeldingen zijn deze items niet geselecteerd; 2) Er is tenminste één spiegelitem dat hetzelfde leerdoel meet met een vergelijkbare vraagstelling (zie Bijlage 1); 3) de items hadden voldoende *face-validity*; 4) de items werden door twee expertleerkrachten van groep 7 beoordeeld als voldoende discriminerend. De toets werd in de digitale leeromgeving van Cito afgenomen. 13 items bevatten een relevante afbeeldingen met *seductive details*, de andere 13 een relevante afbeelding zonder *seductive details*. In tabel 2 is te zien hoe de originele en aangepaste items zijn verdeeld over de twee toetsversies.

Tabel 2

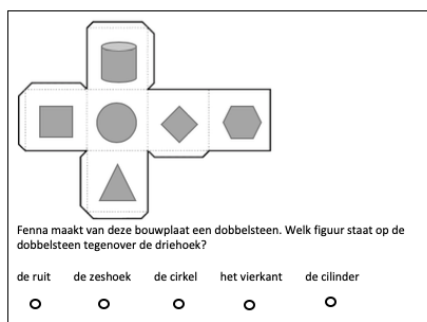
Verdeling van de Originele en de Aangepaste Toetsitems over de Twee Toetsversies

	Toetsitems origineel	Toetsitems aangepast
Toetsversie 1	2; 4; 6; 8; 10; 12; 13; 15; 17; 19; 21; 23; 26	1; 3; 5; 7; 9; 11; 14; 16; 18; 20; 22; 24; 25
Toetsversie 2	1; 3; 5; 7; 9; 11; 14; 16; 18; 20; 22; 24; 25	2; 4; 6; 8; 10; 12; 13; 15; 17; 19; 21; 23; 26

De oorspronkelijke items van de Groeimeter (Cito, 2019) waren de originele items met relevante afbeeldingen met *seductive details*. Om dezelfde stijl te waarborgen van zowel de afbeeldingen in de aangepaste items als in de oorspronkelijke items, werden deze aangepast door dezelfde illustrator. Voor het herontwerp werden de volgende richtlijnen gehanteerd: 1) gekleurde afbeeldingen werden vervangen door zwart-wit afbeeldingen; 2) alle overbodige details, zoals menselijke kenmerken, werden geëlimineerd; 3) authentieke elementen werden vervangen door meer abstracte elementen, bijv. een cirkel in plaats van een appeltaart. Een voorbeeld van de gemaakte aanpassingen is te zien in Figuur 1a en 1b (Cito, 2019).



Figuur 1a. Voorbeelditem: relevante afbeelding met *seductive details* uit de Groeimeter. Copyright 2019 by Cito.



Figuur 1b. Voorbeelditem: relevante afbeelding zonder *seductive details* uit de Groeimeter. Copyright 2019 by Cito.

Voor het maken van de toets gold geen tijdslimiet, omdat dit in de oorspronkelijke versie ook geen rol speelt. Wanneer tijdslimiet wordt ingevoerd, terwijl dit normaal niet het geval is, kan deze tijdsdruk een vertekend beeld van het effect geven (Ögren et al., 2017). De tijd, die de leerling besteedde aan het beantwoorden van het item, werden door de toetssoftware wel geregistreerd in seconden. Hierdoor kon bij de analyse worden bekeken of de benodigde tijd voor het beantwoorden van de originele en aangepaste items verschilde, aangezien Beddow (2018) aangeeft dat *seductive details* onnodig veel tijd kosten, die ten koste gaat van de tijd voor het beantwoorden van het item.

3.3.2 Meten van interesse

Het is van belang dat de interesse op een effectieve en efficiënte manier gemeten wordt, zodat het niet te veel cognitieve capaciteit kost. In een eerder onderzoek van Lindner et al. (2016) naar de interesse bij een test betreffende natuurwetenschappen is succesvol gebruikt gemaakt van een 4-punts beoordelingsschaal. Daarbij werd op een eenvoudige manier gevraagd hoeveel interesse de lerende had voor ieder item. In huidig onderzoek werd een soortgelijke beoordelingsschaal gebruikt. Na ieder toetsitem beoordeelde de lerende hoeveel interesse er was voor het toetsitem door de vraag een cijfer te geven van 1 (“helemaal niet leuk”) tot 4 (“erg leuk”). Door *emojicons* te gebruiken werd deze getalschaal verduidelijkt (Lindner et al., 2016, p. 4.). De tekst behorend bij het cijfer kwam in beeld op het moment dat op het cijfer werd geklikt of het blauwe vierkantje naar dit cijfer werd verschoven. Op basis van Lindner et al. (2016) werd voor huidig onderzoek een eigen versie gemaakt (zie Figuur 2).



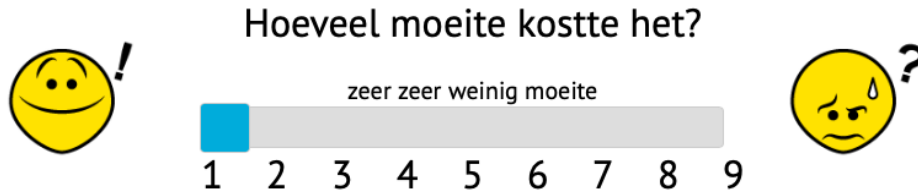
Figuur 2. Beoordelingsschaal interesse gebruikt in huidig onderzoek. Aangepast van “How representational pictures enhance students’ performance and test-taking pleasure in low-stakes assessment” van Lindner, M. A., Ihme, J. M., Saß, S. & Köller, O., 2016, *European Journal of Psychological Assessment*, p. 4. Copyright 2016 by Hogrefe Publishing.

3.3.3 Meten van de geleverde cognitieve belasting

In eerder onderzoek naar de multimedia principes is vaak de geleverde cognitieve belasting gemeten door een zelfbeoordeling van de lerende te vragen. In het huidige onderzoek werd de 9-punts beoordelingsschaal voor cognitieve belasting van Paas (1992) gebruikt. Het is een veel gebruikte beoordelingsschaal om cognitieve belasting te meten. Verschillende onderzoeken geven aan dat deze beoordelingsschaal geschikt is voor het meten van cognitieve belasting en dat het geldige informatie geeft over de individuele ervaring (Gopher & Braune, 1984; Paas & Van Merriënboer, 1994). In meer recent onderzoek (Korbach, Brinken, & Park, 2018; Szulewski, Gegenfurtner, Howes, Sivilotti, & Van Merriënboer, 2018), waarin de verschillende meetinstrumenten voor cognitieve belasting met elkaar worden vergeleken, blijkt tevens dat deelnemers de individuele taakmoeilijkheid op betrouwbare wijze kunnen beoordelen met de gebruikte beoordelingsschaal.

Daarnaast kan de beoordelingsschaal afgenomen worden na het afronden van elke taak (Van Gog, Kirschner, Kester, & Paas, 2012). Dit is een voordeel ten opzichte van andere meetinstrumenten die cognitieve belasting meten aan het eind van een reeks taken. Dit type meting vereist dat leerlingen een retrospectieve beoordeling geven van de cognitieve belasting en dat moet waarschijnlijk op worden gehaald uit het lange termijn geheugen, waardoor een vertekening plaats vindt van de zelfinschatting

(Van Gog et al., 2012). Tevens is gebleken dat de beoordelingsschaal van Paas gebruiksvriendelijk is voor leerlingen in het primair onderwijs wanneer door de toevoeging van *emoticons* de getalschaal wordt verduidelijkt (zie Figuur 3; Gordon, Tindall-Ford, Agostinho, & Paas, 2016).



Figuur 3. Beoordelingsschaal cognitieve belasting gebruikt in huidig onderzoek. Aangepast van “Learning from Instructor-managed and Self-managed Split-attention Materials” van Gordon, C., Tindall-Ford, S., Agostinho, S., & Paas, F., 2016, *Applied Cognitive Psychology*, 30(1), p. 5. Copyright 2015 by John Wiley & Sons, Ltd.

Na ieder toetsitem beoordeelde de leerling de cognitieve belasting door het item een cijfer te geven van 1 (‘zeer zeer weinig moeite’) tot 9 (‘zeer zeer veel moeite’). Ook bij deze schaal verscheen de bijbehorende tekst bij het cijfer op het moment dat de leerling op het cijfer klikte of het blauwe vierkantje naar dit cijfer verschoof.

3.4 Procedure

Het onderzoek werd goedgekeurd door de ethische commissie van de Open Universiteit op 10 mei 2019. Voor het werven van de participanten voor het onderzoek werd de directeur van een school voor primair onderwijs benaderd. Vervolgens werden de leerkrachten van groep 7 van deze school door de directeur benaderd. Het doel en de opzet van het onderzoek werden schriftelijk en mondeling toegelicht. Na schriftelijke toestemming van de directie werden de ouders/verzorgers van alle groep 7 leerlingen schriftelijk benaderd via de groepsleerkracht. In de brief werden de ouders/verzorgers middels een *passive informed consent* geïnformeerd over het doel van het onderzoek en waarom dit onderzoek bij zou kunnen dragen aan een verbetering van de kwaliteit van de digitale toetsen. Wanneer ouders/verzorgers niet wilden dat hun kind deelnam moesten ze dit aangeven bij de groepsleerkracht of de onderzoeker. Dit was het geval bij drie leerlingen. Zij bleven tijdens de afname in het lokaal en maakten niet aan de toets gerelateerd werk dat ze ontvingen van de leerkracht. In de brief naar de ouders werd verder toegelicht wat de onderzoeksprocedure was, alsmede de duur. Er werd benadrukt dat deelname geheel vrijwillig was en dat ouders op ieder moment het recht van deelname van hun kind konden beëindigen. De anonimiteit van hun kind werd gegarandeerd en de ouders konden, middels de in de brief opgenomen contactinformatie, contact opnemen met de onderzoekers.

De toets werd per groep op één ochtend afgenomen in het eigen klaslokaal van de leerlingen door de onderzoeker, de eigen leerkracht was hierbij aanwezig. De tafels stonden in een zogenaamde toetsopstelling, waardoor bij elkaar kijken bemoeilijkt werd. De leerlingen kregen een mondelinge uitleg over het doel en de opzet van het onderzoek, gevolgd door een uitleg over hoe de online omgeving werkte, welke werd gevisualiseerd op het digitale schoolbord. Aan de leerlingen werd niet verteld welke aanpassingen zijn gedaan aan de originele toetsitems. Bij de instructie hoorden twee voorbeelditems, die zijn voorgedaan door de onderzoeker op het digitale schoolbord. Bij de voorbeelditems werden beide beoordelingsschalen, die na ieder item volgden, toegelicht. Daarbij werd benadrukt dat het antwoord op deze vragen niet goed/fout kon zijn, omdat werd gevraagd naar de mening van de leerlingen. Er werd tevens verteld dat terugkeren naar eerdere items niet mogelijk was.

Vervolgens maakten de leerlingen in stilte de toets in hun eigen tempo op een laptop van de school waar werd ingelogd in de online toets omgeving van Cito. Na ieder toetsitem werd de leerlingen gevraagd om aan te geven hoe leuk zij het toetsitem vonden en hoeveel cognitieve belasting zij hadden geleverd bij het maken van de opgave. Deze vragen verschenen vanzelf na ieder toetsitem en het juiste antwoord kon worden aangeklikt. De leerlingen konden pas doorklikken naar het volgende toetsitem als de vragen over de interesse en de cognitieve belasting waren beantwoord. Wanneer de leerlingen klaar waren bleven zij in de ruimte werken aan een zelfstandige opdracht, waardoor anderen niet werden gestoord. Deze opdracht had geen relatie met het onderzoek. Als dank voor deelname aan het onderzoek kregen de deelnemende groepen 7 elk een cadeaubon om spelmateriaal of boeken aan te kunnen schaffen.

3.5 Data-analyse

Na afloop van de toets werd de data met de gegevens van alle benodigde variabelen bij Cito opgevraagd. Deze data werd geëxporteerd naar het programma SPSS (versie 24). Er werd eerst een vooranalyse gedaan om te bekijken of er bepaalde items heel makkelijk of moeilijk waren. Bij één item, dat vele kinderen fout hebben beantwoord, werd de sleutel kritisch bekeken, deze bleek kloppend te zijn en er is besloten dit item in de analyse te laten. Vervolgens werd een onafhankelijke t -toets gedaan om de toetsresultaten per toetsversie te analyseren. Dit werd gedaan om er zeker van te zijn dat beide toetsversies gelijk waren qua moeilijkheidsgraad. Tijdens deze vooranalyse werd tevens met een onafhankelijke t -toets bekeken of er verschillen waren in interesse, cognitieve belasting, toetsresultaat en benodigde tijd tussen jongens en meisjes. Daardoor kon bekeken worden of eventuele verschillen toegeschreven konden worden aan de verschillen in geslacht.

Vervolgens bestond de analyse van de data uit vier delen, aansluitend bij de opgestelde hypotheses. Allereerst werd met een gepaarde t -toets onderzocht of de interesse bij de aangepaste items zonder *seductive details* verschilde met die van de originele items met *seductive details*. Dit werd tevens met een post-hoc analyse bekeken per leerdoel met een gepaarde t -toets. In het tweede deel van de analyse werd er een gepaarde t -toets uitgevoerd om te onderzoeken of de cognitieve belasting bij de

aangepaste items zonder *seductive details* lager zou zijn dan bij de originele items met *seductive details*. Ook dit werd vervolgens post-hoc bekeken per leerdoel met een gepaarde *t*-toets. Het toetsresultaat werd tevens onderzocht met een gepaarde *t*-toets om te onderzoeken of de aangepaste items zonder *seductive details* beter waren gemaakt dan de originele items met *seductive details*, waarna ook hier met dezelfde post-hoc analyse gekeken werd of er verschillen waren in het toetsresultaat per leerdoel. Tenslotte werd met een gepaarde *t*-toets geanalyseerd of er verschillen waren tussen de originele en aangepaste items met betrekking tot de benodigde tijd om het item te beantwoorden en vervolgens in de post-hoc analyse ook per leerdoel. Bij alle analyses van de hypothesen werd uitgegaan van een alpha level van 0.05 en een betrouwbaarheidsinterval van 95%.

4. Resultaten

In dit onderzoek ging het om de vraag of de interesse, cognitieve belasting, het toetsresultaat en de benodigde tijd veranderen naarmate de gebruikte afbeeldingen in de toets zijn aangepast volgens het coherentieprincipe. In tabel 3 zijn de gemiddelden en standaarddeviaties van de berekende variabelen gepresenteerd.

Tabel 3

Gemiddelden en Standaard Deviaties van Getoetste Variabelen

	<i>n</i>	Origineel		Aangepast	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Interesse	67	2.66	0.57	2.46	0.61
Cognitieve belasting	67	2.76	1.00	2.79	1.05
Toetsresultaat	67	.81	.14	.81	.12
Benodigde tijd	67	57.11	15.12	55.64	13.91

4.1 Vooranalyse

Er werd gekeken of de moeilijkheidsgraad per toetsversie verschilde. Toetsversie 2 ($M = 21.59$, $SD = 2.54$) is beter gemaakt dan toetsversie 1 ($M = 20.60$, $SD = 2.99$), dit verschil is echter niet significant, $t(65) = -1.46$, $p = .15$. Tevens werd geanalyseerd of er een verschil in toetsresultaat was tussen de jongens en de meisjes. Jongens hebben de toets beter gemaakt ($M = 21.50$, $SD = 2.71$) dan meisjes ($M = 20.73$, $SD = 2.87$). Dit verschil was niet significant, $t(65) = 1.12$, $p = .27$. Ook het verschil in interesse tussen jongens ($M = 2.42$, $SD = 0.58$) en meisjes ($M = 2.67$, $SD = 0.51$) bleek tijdens de vooranalyse niet significant, $t(65) = -1.90$, $p = .06$. Bij de vooranalyse van de cognitieve belasting bleken meisjes ($M = 3.18$, $SD = 0.99$) in het algemeen een hogere cognitieve belasting te ervaren dan de jongens ($M = 2.27$, $SD = 0.83$). Dit verschil was wel significant, $t(65) = -4.01$, $p = <.001$. Meisjes bleken ook meer tijd nodig te hebben voor het beantwoorden van de items ($M = 58.28$, $SD = 13.36$), dan de jongens ($M = 54.02$, $SD = 12.75$). Dit verschil was echter niet

significant, $t(65) = -1.32$, $p = .19$. Doordat de meeste variabelen niet significant verschilden tussen jongens en meisjes is er besloten om de data van de jongens en de meisjes en de twee toetsversies samen te voegen en de analyses over de hele dataset uit te voeren.

4.2 Interesse

Hypothese één veronderstelde dat leerlingen meer interesse hebben voor de items met relevante afbeeldingen met *seductive details* dan voor de items met relevante afbeeldingen zonder *seductive details*. Om te toetsen of dit inderdaad het geval was, werd er een gepaarde *t*-toets uitgevoerd met de *within-subject* factor ‘type afbeelding’ en de afhankelijke variabele ‘interesse’. De originele items met *seductive details* ($M = 2.66$, $SD = 0.57$) werden interessanter bevonden dan de aangepaste items zonder *seductive details* ($M = 2.55$, $SD = 0.60$). Dit verschil was significant, $t(66) = -4.27$, $p = <.001$, en bevestigde daarmee de eerste hypothese.

4.3 Cognitieve belasting

Om te toetsen of er een verschil was tussen de cognitieve belasting bij de toetsitems zonder *seductive details* en de toetsitems met *seductive details* werd er een gepaarde *t*-toets uitgevoerd met de *within-subject* factor ‘type afbeelding’ en de afhankelijke variabele ‘cognitieve belasting’. De tweede hypothese stelde dat de leerlingen bij de toetsitems zonder *seductive details* een lagere cognitieve belasting zouden rapporteren dan bij de toetsitems met *seductive details*. Het verschil in de cognitieve belasting voor de aangepaste items ($M = 2.79$, $SD = 1.05$) en de originele items ($M = 2.76$, $SD = 1.04$) was echter niet significant, $t(66) = .45$, $p = .65$. De tweede hypothese werd daarmee verworpen.

4.4 Toetsresultaat

De derde hypothese stelde dat er een verschil was in toetsresultaat tussen de toetsitems zonder *seductive details* en de toetsitems met *seductive detail*. Leerlingen zouden op basis van de theorie over het coherentieprincipe de items zonder *seductive details* beter maken dan de originele items met relevante afbeeldingen met *seductive details*. Dit werd getoetst met een gepaarde *t*-toets met de *within-subject* factor ‘type afbeelding’ en de afhankelijke variabele ‘toetsresultaat’. De aangepaste items ($M = .81$, $SD = .12$) werden niet beter gemaakt dan de originele items ($M = .81$, $SD = .14$). Dit verschil was dan ook niet significant, $t(66) = .26$, $p = .79$. Deze resultaten zorgen ervoor dat ook de derde hypothese werd verworpen.

4.5 Benodigde tijd

De laatste hypothese stelde dat leerlingen meer tijd nodig zouden hebben om de toetsitems met *seductive details* te beantwoorden dan de toetsitems zonder *seductive detail*. Dit bleek niet het geval te zijn, de originele items (met *seductive details*) kostten weliswaar iets meer tijd om te beantwoorden ($M = 57.11$, $SD = 15.13$) dan de aangepaste items ($M = 55.64$, $SD = 13.91$), maar dit verschil was niet significant, $t(66) = -.98$, $p = .33$. Ook de laatste hypothese werd daardoor verworpen.

4.6 Post-hoc analyses

Aangezien de items per leerdoel een verschillende vraagstelling hadden zijn er post-hoc analyses uitgevoerd voor de afhankelijke variabelen van een aantal leerdoelen. De resultaten daarvan zullen hieronder besproken worden. Bij vijf leerdoelen was het verschil in interesse tussen het originele en het aangepaste item significant. De leerlingen vonden daarbij steeds het originele item met *seductive details* leuker dan het aangepaste item zonder *seductive details* (zie Tabel 4).

Tabel 4

Gemiddelden en Standaard Deviaties van Items met Significant Verschil op Interesse

	Cognitieve belasting					
	<i>n</i>	Origineel		Aangepast		<i>p</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
lego	67	2.76	0.95	2.28	0.98	.01
tellen	67	2.46	0.96	2.18	1.09	.03
dobbelsteen	67	2.84	0.88	2.30	1.04	<.001
reeksen	67	3.06	1.01	2.61	1.07	.002
meten	67	2.34	1.04	2.01	0.95	.04

Bij verdere analyse werden de cognitieve belasting, het toetsresultaat en de benodigde tijd van bovenstaande leerdoelen waarbij de interesse significant verschilde, nader bekeken. Bij geen van deze leerdoelen was het verschil in cognitieve belasting of toetsresultaat significant (zie Bijlage 2 en 3). Opvallend was wel dat, behalve bij het leerdoel ‘meten’, de originele items met de *seductive details* minder cognitieve belasting kostten dan de aangepaste items en dat de originele items beter werden gemaakt dan de aangepaste items. Voor de factor tijd bleken alleen de items waarbij de leerlingen een reeks moesten afmaken, significant te verschillen, waarbij opvallend genoeg het aangepaste item zonder *seductive details* meer tijd kostte om te beantwoorden dan het originele item (zie Bijlage 4).

5. Conclusie en discussie

Het doel van dit onderzoek was om het effect van *seductive details* bij relevante afbeeldingen in een rekentoets te onderzoeken. Vanuit de theorie van CTML mocht verondersteld worden dat de toetsitems met relevante afbeeldingen met *seductive details* een negatief effect zouden hebben op de geleverde cognitieve belasting, het toetsresultaat en de benodigde tijd. Echter wanneer vanuit de benadering van interesse wordt gekeken naar het effect van *seductive details* in relevante afbeeldingen, zouden deze details juist een positief effect op het resultaat kunnen hebben, omdat deze afbeeldingen kunnen zorgen voor een verhoogde interesse.

5.1 Conclusie

De leerlingen waren in huidig onderzoek meer geïnteresseerd in de toetsitems met relevante afbeeldingen met *seductive details*. Deze waren dus niet aangepast volgens het coherentieprincipe. De toetsitem waarbij de *seductive details* werden geëlimineerd in de relevante afbeeldingen zorgden niet voor een verlaging van de geleverde cognitieve belasting, een verbetering van het toetsresultaat of een vermindering van de benodigde tijd. Op basis van de resultaten in huidig onderzoek lijkt het voor digitale toetsen dus niet zinvol om het coherentieprincipe toe te passen wanneer er relevante afbeeldingen worden gebruikt. Het tegenovergestelde kan zelfs geconcludeerd worden; het verdient aanbeveling om *seductive details* aan de relevante afbeeldingen in een toets toe te voegen, omdat dit zorgt voor een omgeving waarin leerlingen meer geïnteresseerd zijn, zonder dat het de cognitieve belasting verhoogt of de resultaten en benodigde tijd negatief beïnvloedt.

5.2 Discussie

5.2.1 Interesse

Het enige significante resultaat van huidig onderzoek is dat leerlingen meer geïnteresseerd zijn in de toetsitems met relevante afbeeldingen met *seductive details* dan in items zonder *seductive details*. Dit bevestigt eerder onderzoek over het ontwerp van lesmateriaal, waarbij is gesteld dat het toevoegen van menselijke kenmerken en verschillende aantrekkelijke kleuren factoren zijn, die een les aantrekkelijker kunnen maken (Mayer & Estrella, 2014). Dit bleek in huidig onderzoek dus ook voor (formatief) toetsmateriaal het geval te zijn.

In experimenten van Mayer et al. (2008) blijkt dat naarmate de lerende de details in de relevante afbeeldingen interessanter vonden, het begrip afnam. Echter in huidig onderzoek bij toetsen werd dit niet bevestigd. Bij vrijwel alle leerdoelen, waarbij de interesse significant verschilde, werd het originele item met de *seductive details* beter gemaakt dan het aangepaste item zonder *seductive details*, ook al was dit verschil niet significant.

Voor vervolgonderzoek naar het effect van het toevoegen van menselijke kenmerken en kleuren op toetsen kan er onderzocht worden of bij toetsen vooral de menselijke kenmerken of het kleurgebruik het meest interessant worden bevonden. Volgens de CATLM (Moreno, 2005) kan ontwerp, waarbij menselijke kenmerken en kleuren zijn toegevoegd, zorgen voor een actievere verwerking van de gepresenteerde informatie. Leerlingen zouden door de toevoegingen van de menselijke kenmerken en de kleuren hun aandacht beter kunnen richten op de relevante informatie (i.e. selecteren), dit wordt door onderzoek van Izard en Ackerman (2000) en Park et al. (2015) ondersteund, en het zou leerlingen kunnen aanmoedigen om het relevante materiaal beter te organiseren en te integreren. Het is niet duidelijk of deze voordelen optreden door het toevoegen van menselijke kenmerken of kleuren.

5.2.2 Cognitieve belasting

Er werd geen verschil gevonden in zelfgerapporteerde cognitieve belasting tussen de toetsitems met relevante afbeeldingen zonder *seductive details* en de toetsitems met relevante afbeeldingen met *seductive details*. Dit is in strijd met het onderzoek van Sung en Mayer (2012) en Wang en Adesope (2014), die stellen dat *extraneous cognitive load* wordt gecreëerd door de *seductive details*, omdat deze afleiden van de essentiële informatie. Een verklaring kan liggen in het feit dat aanwezige voorkennis een bepalend kenmerk van de lerende blijkt te zijn voor de invloed van het effect van multimediacprincipes voor leren (Kalyuga, Chandler, & Sweller, 1998). Dit sluit aan bij de CLT waarbij wordt gesteld dat *intrinsic cognitive load* verlaagd kan worden door domeinspecifieke kennis en dat deze belasting lager is voor lerenden met meer voorkennis dan voor lerenden met minder voorkennis. Dus het effect van *seductive details* zou leerlingen met een lagere voorkennis meer beïnvloeden dan leerlingen met een hogere voorkennis (Park, Korbach, & Brünken, 2015; Magner, Schwonke, Aleven, Popescu, & Renkl, 2014). Vertaald naar toetsen zou men kunnen veronderstellen dat leerlingen die beter presteren op de gehele toets (en dus meer kennis hebben over het onderwerp) minder baat hebben bij het aanpassen van de items volgens het coherentieprincipe dan leerlingen die minder goed presteren op de toets. Het is interessant om bij vervolgonderzoek aanwezige voorkennis of algehele toetsresultaat bij de leerling als variabele mee te nemen en te onderzoeken of hier een verklaring ligt voor het feit dat *seductive details* in huidig onderzoek geen nadelig effect hadden op de geleverde cognitieve belasting bij toetsen.

Een alternatieve verklaring voor de gevonden resultaten zou kunnen zijn dat het gebruikte meetinstrument om de cognitieve belasting te meten niet voldoende valide was aangezien de leerlingen een zelfbeoordeling moesten geven en er geen onderscheid gemaakt kon worden in de verschillende soorten cognitieve belasting (Sweller et al., 2019). Voor vervolgonderzoek kan het daarom nuttig zijn een meer objectief meetinstrument in te zetten om cognitieve belasting te meten (Korbach et al., 2018), dan de in huidig onderzoek gebruikte meer subjectieve beoordelingsschaal van Paas (1992). Een voorbeeld van een effectief instrument is *eye-tracking* (Korbach et al., 2018), omdat daarmee onverwachte bevindingen verklaard kunnen worden (Jarodzka, Holmqvist, & Gruber, 2017). Door *eye-tracking* te gebruiken kan worden onderzocht wat de totale fixatietijd is op relevante informatie (Korbach et al., 2018) en kan geanalyseerd worden wanneer de leerling voor het eerst naar de relevante informatie kijkt (Park et al., 2015). De eerste fixaties van de leerlingen zouden de meeste informatie geven over de invloed van *seductive details*. Het meten van de fixatie kan zeker voor toetsen relevant zijn, omdat het doel van toetsen kan zijn om te beoordelen of het de lerende lukt om in een toetssituatie relevante informatie te filteren van irrelevante informatie (Jarodzka et al., 2015). Meer kennis verbetert de informatieselectie. Tevens kan *eye-tracking* inzicht geven in de cognitieve processen van selecteren, organiseren en integreren, zodat de actieve verwerking geanalyseerd kan worden (Alemdag & Cagiltay, 2018). Er kan zelfs een verband worden gelegd met de interesse, omdat

uit eerder onderzoek blijkt dat onderzoekers door *eye tracking* conclusies kunnen trekken over welke items interessant waren voor de participanten en hoe deze werden waargenomen (Duchowski, 2003).

5.2.3 Toetsresultaat en tijd

Wanneer we de eerdergenoemde gevonden resultaten bij de cognitieve belasting vergelijken met de in huidig onderzoek gevonden resultaten bij de variabele ‘toetsresultaat’ dan zijn deze passend bij de CTML. Het elimineren van de *seductive details* zou moeten leiden tot een verlaging van de cognitieve belasting, waardoor de resultaten zouden verbeteren. Aangezien in huidig onderzoek echter geen effect is gevonden op de cognitieve belasting is het vanuit de CTML gedachte logisch dat dit effect ook niet gevonden is bij de toetsresultaten. Dat er geen gunstig effect is gevonden op de toetsresultaten komt echter niet overeen met onderzoek naar *seductive details* van onder andere Butcher (2006), Mayer en Fiorella (2014) en Wang en Adesope (2014). In deze onderzoeken presteert de groep lerenden met afbeeldingen zonder *seductive details* beter dan de groep die de afbeeldingen met *seductive details* bekeken. Het betreft hier echter allen situaties waarin werd geleerd, terwijl huidig onderzoek ging over toetsen.

Wanneer we de toetsresultaten bekijken vanuit de theorie dat een toename van interesse zorgt voor betere resultaten, hadden we mogen verwachten dat de toetsresultaten, van de items waarbij de interesse significant verschilde, een verschil lieten zien. Een hogere interesse zou gezorgd kunnen hebben voor een hogere inspanning (Izard & Ackerman, 2000) en een volledige inzet van de cognitieve middelen (Moreno, 2005) waardoor de resultaten beter zouden moeten zijn. In huidig onderzoek werd dit effect echter niet gevonden. Dit sluit aan bij Beddow (2018) die stelt dat *seductive details* geen verschil in toetsresultaat opleveren. Zijn bevinding dat de items met *seductive details* onnodig meer tijd zouden kosten dan de items zonder deze details werd echter ook niet bevestigd in huidig onderzoek. Bij het enige leerdoel waarbij de benodigde tijd significante verschilde, bleek zelfs het tegendeel waar; het aangepaste, vereenvoudigde item kostte meer tijd dan het originele item.

Een verklaring zou kunnen liggen in het ruimtelijk inzicht van leerlingen. Voor leerlingen met een groter ruimtelijk inzicht is het makkelijker om mentale modellen te construeren uit gegeven informatie in zowel de tekst als in de afbeeldingen (Meneghetti, Gyselinck, Pazzaglia, & De Beni, 2009). Tevens blijkt dat deze leerlingen bij leren minder beïnvloed werden door *seductive details* (Park et al., 2015), dit effect bleek het grootste voor informatie in afbeeldingen. Echter bij toetsen kan dit effect minder zichtbaar zijn, aangezien de benodigde kennis en vaardigheden met betrekking tot het ruimtelijk inzicht reeds zijn verworven. Dit is een mogelijke verklaring waarom er geen verschil in toetsresultaat en benodigde tijd is gevonden tussen toetsitems zonder *seductive details* en toetsitems met *seductive details*. Voor vervolgonderzoek kan het interessant zijn om het ruimtelijk inzicht mee te nemen.

5.3 Beperkingen van het onderzoek en aanbevelingen voor de praktijk

De steekproef van dit onderzoek bestond uit 67 kinderen van één school. Voor vervolgonderzoek verdient het aanbeveling om een grotere groep leerlingen te onderzoeken van verschillende scholen. Verder was de onderzoeksgroep niet random getrokken, waardoor de steekproef geen afspiegeling is voor de populatie basisschoolkinderen in groep 7 in Nederland. Er werd bij de leerlingen niet gekeken naar individuele verschillen in bijvoorbeeld voorkennis en ruimtelijk inzicht, terwijl dit wel invloed kan hebben op de resultaten. Het kan waardevol zijn om dit in volgend onderzoek mee te nemen.

De items zijn oorspronkelijk ontworpen voor groep 6, maar doordat het onderzoek in de eerste periode van het schooljaar plaatsvond, is deze in groep 7 afgenomen. Voor een vervolgonderzoek kan het zinvol zijn om de toetsitems af te nemen in groep 6, zodat de toetsdoelen beter passend zijn bij de aangeboden leerstof en de aanwezige kennis lager is bij de leerlingen. Het aantal onderzochte items was verder beperkt, met een grote variatie aan leerdoelen. Ook dit zou bij vervolgonderzoek een aandachtspunt kunnen zijn. Wanneer dan tevens de cognitieve belasting gemeten wordt met een meer objectief beoordelingsinstrument dan met de in huidig onderzoek gebruikte beoordelingsschaal van Paas (1992) zullen de resultaten nog waardevoller kunnen zijn voor het maken van richtlijnen voor digitale toetsen.

Ondanks deze beperkingen geeft dit onderzoek zicht op het effect van *seductive details* in toetsitems. Doordat in huidig onderzoek zelfs een negatief effect gevonden is voor het toepassen van het coherentieprincipe voor de interesse kan het waardevol zijn voor ontwerpers van digitale toetsen, zoals Cito, om *seductive details* in relevante afbeeldingen te blijven gebruiken om daarmee een omgeving te creëren die prettiger oogt voor de leerlingen. De aandacht wordt daardoor gevestigd op de inhoud en wellicht dat daardoor de inspanning van de lerende wordt vergroot.

6. Referenties

- Alemdag, E. & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computer & Education*, 125, 413-428. doi:10.1016/j.compedu.2018.06.023
- Baddeley, A. D. (1999). *Human memory*. Boston: Allyn & Bacon.
- Beddow, P. A. (2018). Cognitive Load Theory for test design. In S.N. Elliot, R.J Kettler, P.A. Beddow, & A. Kurz, *Handbook of accessible instruction and testing practices: issues, innovations, and applications* (2nd ed.;pp 199-211). Cham: Springer International Publishing.
- Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, 98, 182-197. doi:10.1037/0022-0663.98.1.182
- Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.). *The Cambridge handbook of multimedia learning* (pp. 174–205). (2nd ed.). New York, NY: Cambridge University Press.
- Cito (2019). *Groeimeter*. Geraadpleegd van <http://afname.citolab.nl>
- Clark, R. C., & Mayer, R. E. (2016). *e-Learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning*, (Fourth Edition ed.). New Jersey: John Wiley & Sons.
- Creswell, J. W. (2014). *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Essex: Pearson.
- De Westelinck, K., Valcke, M., De Craene, B., & Kirschner, P. A. (2005). Multimedia learning in social sciences: Limitations of external graphical representations. *Computers in Human Behavior*, 21(4), 555-573. doi:10.1016/j.chb.2004.10.030
- Duchowski, A. T. (2003). *Eye tracking methodology: Theory and practice* (2nd ed.). London: SpringerVerlag.
- Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 26(5), 519-532. doi: 10.1177/001872088402600504
- Gordon, C., Tindall-Ford, S., Agostinho, S., & Paas, F. (2016). Learning from Instructor-managed and self-managed split-attention materials. *Applied Cognitive Psychology*, 30(1), 1-9. doi:10.1002/acp.3159
- Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: a theory of cognitive interest in science learning. *Journal of Educational Psychology*, 90, 414-434. doi:10.1037/0022-0663.90.3.414
- Izard, C. E., & Ackermann, B. P. (2000). Motivational, organizational, and regulatory functions of discrete emotions. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of emotion*, Vol. II (pp. 253-264). New York: Guilford.

- Jarodzka, H., Holmqvist, K., & Gruber, H. (2017). Eye tracking in Educational Science: Theoretical frameworks and research agendas. *Journal of Eye Movement Research*, 10(1): 3,1-18
doi:10.16910/jemr.10.1.3
- Jarodzka, H., Janssen, N., Kirschner, P. A., & Erkens, G. (2015). Avoiding split attention in computer-based testing: Is neglecting additional information facilitative? *British Journal of Educational Technology*, 46, 803-817. doi:10.1111/bjet.12174
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23–31. doi:10.1207/S15326985EP3801_4
- Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40, 1–17.
doi:10.1518/001872098779480587
- Kirschner, P. A., Park, B., Malone, S., & Jarodzka, H. (2016). Toward a cognitive theory of multimedia assessment (CTMMA). In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, design, and technology: An international compendium of theory, research, practice, and policy* (pp. 1-23). Cham: Springer International Publishing.
- Korbach, A., Brünken, R. & Park, B. (2018) Differentiating different types of cognitive load: a comparison of different measures. *Educational Psychology Review*, 30, 503–529.
doi:10.1007/s10648-017-9404-8
- Lindner, M. A., Eitel, A., Strobel, B., & Köller, O. (2017a). Identifying processes underlying the multimedia effect in testing: An eye-movement analysis. *Learning and Instruction*, 47, 91-102.
doi:10.1016/j.learninstruc.2016.10.007
- Lindner, M. A., Ihme, J. M., Saß, S., & Köller, O. (2016). How representational pictures enhance students' performance and test-taking pleasure in low-stakes assessment. *European Journal of Psychological Assessment*, 34, 376-385. doi:10.1027/1015-5759/a000351
- Lindner, M. A., Lüdtke, O., Grund, S., & Köller, O. (2017b). The merits of representational pictures in educational assessment: Evidence for cognitive and motivational effects in a time-on-task analysis. *Contemporary Educational Psychology*, 51, 482-492. doi:10.1016/j.cedpsych.2017.09.009
- Magner, U. I. E., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction*, 29, 141-152. doi:10.1016/j.learninstruc.2012.07.002
- Manastirean-Zijlstra, C. S. (2017). *Minder zoeken, beter toetsen. Een eye tracking onderzoek naar de toepassing van de instructieprincipes uit de Cognitive Theory of Multimida Learning bij het ontwerpen van computergebaseerde toetsing [master scriptie]*. Heerlen, Open Universiteit.
- Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning*. (2nd ed.). New York: Cambridge University Press.
- Mayer, R. E., & Estrella, E. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction*, 33, 12-18. doi:10.1016/j.learninstruc.2014.02.004

- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity and, temporal contiguity principles. In R. E. Mayer (Ed.). *The Cambridge handbook of multimedia learning* (pp. 279–315). (2nd ed.). New York, NY: Cambridge University Press.
- Mayer, R. E., Griffith, E., Jurkowitz, I. T. N., & Rothman, D. (2008). Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied*, 14(4), 329–339. doi:10.1037/a0013835
- Meneghetti, C., Gyselinck, V., Pazzaglia, F., & De Beni, R. (2009). Individual differences in spatial text processing: High spatial ability can compensate for spatial working memory interference. *Learning and Individual Differences*, 19, 577–589. doi:10.1016/j.lindif.2009.07.007
- Moreno, R. (2005). Instructional technology: Promise and pitfalls. In L. PytlikZillig, M. Bodvarsson & R. Bruning (Eds.), *Technology-based education: Bringing researchers and practitioners together* (pp 1-19). Greenwich, CT: Information Age Publishing.
- Ögren, M., Nyström M., & Jarodzka H. (2017). There's more to the multimedia effect than meets the eye: is seeing pictures believing? *Instructional Science*, 45(2), 263–287. doi:10.1007/s11251-016-9397-6
- Paas, F. G. W. C. (1992). Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach. *Journal of Educational Psychology*, 84(4), 429–434. doi:10.1037/0022-0663.84.4.429
- Paas, F., & Van Merriënboer, J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6(4), 351–371. doi:10.1007/bf02213420
- Park, B., Flowerday, T., & Brünken, R. (2015). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human Behavior*, 44, 276–278. doi:10.1016/j.chb.2014.10.061
- Park, B., Korbach, A., & Brünken, R. (2015). Do Learner Characteristics Moderate the Seductive-Details-Effect? A Cognitive- Load-Study Using Eye-Tracking. *Educational Technology & Society*, 18(4), 24–36. Geraadpleegd van https://pdfs.semanticscholar.org/1de7/13b70583e89c21d05534ffedf66d1c0bb69f.pdf?_ga=2.225422105.2044687424.1581449721-1323586269.1575967662
- Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.
- Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review*, 7, 216–237. doi:10.1016/j.edurev.2012.05.003
- Sanders, P., Roelofs, E., Visser, J., Wools, S., Meijer, E., Noijons, J., Erkens, T., Kuhlemeier, H., Verstralen, H., Hemker, B. (2017). *Toetsen op school*. Arnhem: Cito.
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.). *The Cambridge handbook of multimedia learning* (pp. 72–103). (2nd ed.). New York, NY: Cambridge University Press.

- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13, 141-156. doi:10.1016/S0959-4752(02)00017-8
- Stárková, T., Lukavsky, J., Javora, O., & Brom, C. (2019). Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning? *Journal of Computer Assisted Learning*, 35, 555-568. doi:10.1111/jcal.12359
- Sung, E., & Mayer, R. E. (2012). When graphics improve liking but not learning from online lessons. *Computers in Human Behavior*, 28(5), 1618-1625. doi:10.1016/j.chb.2012.03.026
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous and germane cognitive load. *Educational Psychology Review*, 22(2), 123-138. doi:10.1007/s10648-010-9128-5
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer. doi:10.1007/978-1-4419-8126-4_6
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional Design. *Educational Psychology Review*, 10(3), 251-296. doi:10.1023/A:1022193728205
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, 31(2), 261-292. doi:10.1007/s10648-019-09465-5
- Szulewski, A., Gegenfurtner, A., Howes, D. W., Sivilotti, M. L. A., & Van Merriënboer, J. J. G. (2018). Measuring physician cognitive load: validity evidence for a physiologic and a psychometric tool. *Advances in Health Sciences Education*, 22(4), 951-968. doi:10.1007/s10459-016-9725-2
- Towler, A., Kraiger, K., Sitzmann, T., Van Overberghe, C., Cruz, J., Ronen, E., & Stewart, D. (2008). The seductive details effect in technology-delivered instruction. *Performance Improvement Quarterly*, 21, 65-86. doi:10.1002/piq.20023
- Valcke, M. (2010). *Onderwijskunde als ontwerpwetenschap: een inleiding voor ontwikkelaars van instructie en voor toekomstige leerkrachten*. Gent: Academia Press.
- Van Gog, T., Kirschner, F., Kester, L., & Paas, F. (2012). Timing and frequency of mental effort measurement: Evidence in favour of repeated measures. *Applied Cognitive Psychology*, 26(6), 833-839. doi:10.1002/acp.2883
- Wang, Z., & Adesope, O. O. (2014). Effects of seductive details on multimedia learning. *Journal of Studies in Education*, 4, 32-44. doi:10.5296/jse.v4i3.6024

Bijlagen

Bijlage 1. Verdeling van de toetsitems

Tabel

Verdeling van de Originele en Aangepaste Toetsitems met de Leerdoelen over de Twee Toetsversies

Naam toetsitem	Leerdoel	Toetsversie 1	Toetsversie 2
Lego bouwwerk	Meten en meetkunde: ruimtelijk inzicht	A ₁ ; O ₁₃	O ₁ ; A ₁₃
Kassabon	Optellen met decimale getallen	O ₂ ; A ₂₂	A ₂ ; O ₂₂
Meten met breuken	Getalbegrip breuken: afmeting object	A ₃ ; O ₁₅	O ₃ ; A ₁₅
Samengestelde breuk	Getalbegrip breuken: samengestelde breuk bij afbeelding	O ₄ ; A ₁₆	A ₄ ; O ₁₆
Staafdiagram tellen	Verbanden: legenda en staafdiagram verbinden	A ₅ ; O ₁₇	O ₅ ; A ₁₇
Dobbelsteen bouwplaat	Meten en meetkunde: wat staat er op de tegenovergestelde zijde?	O ₆ ; A ₁₈	A ₆ ; O ₁₈
Plattegrond	Meten en meetkunde: welke plattegrond hoort bij het bouwwerk?	A ₇ ; O ₁₉	O ₇ ; A ₁₉
Recept	Verhoudingen: welke hoeveelheid heb je nodig voor aangepast recept?	O ₈ ; A ₂₀	A ₈ ; O ₂₀
Reeksen	Meten en meetkunde: maak het patroon af met geometrische vormen	A ₉ ; O ₂₁	O ₉ ; A ₂₁
Prijsbepalen	Decimale getallen: handig rekenen	O ₁₀ ; A ₂₅	A ₁₀ ; O ₂₅
Meten bouwwerk	Opereren met vormen en figuren	A ₁₁ ; O ₂₃	O ₁₁ ; A ₂₃
Staafdiagram	Verbanden: interpreteren van hoeveelheden	O ₁₂ ; A ₂₄	A ₁₂ ; O ₂₄
Optellen	Optellen van hele getallen	A ₁₄ ; O ₂₆	O ₁₄ ; A ₂₆
O = origineel toetsitem A = aangepaste toetsitem			

Bijlage 2. Gemiddelden en standaard deviaties cognitieve belasting bij significant verschil op interesse

Tabel

Gemiddelden en Standaard Deviaties Cognitieve Belasting van Items met Significant Verschil op Interesse

	Cognitieve belasting					
	<i>n</i>	Origineel		Aangepast		<i>p</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
lego	67	3.21	1.87	3.24	1.99	.92
tellen	67	3.42	1.52	3.61	1.65	.34
dobbelsteen	67	2.52	1.64	2.72	1.98	.36
reeksen	67	1.96	1.16	2.03	1.14	.62
meten	67	3.24	2.02	3.16	1.86	.79

Bijlage 3. Gemiddelden en standaard deviaties toetsresultaat bij significant verschil op interesse

Tabel

Gemiddelden en Standaard Deviaties Toetsresultaat van Items met Significant Verschil op Interesse

	<i>n</i>	Toetsresultaat				<i>p</i>
		Origineel		Aangepast		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
lego	67	.57	.50	.55	.50	.88
tellen	67	.75	.44	.72	.45	.71
dobbelsteen	67	.75	.44	.70	.46	.47
reeksen	67	.96	.21	.93	.26	.42
meten	67	.67	.47	.70	.46	.70

Bijlage 4. Gemiddelden en standaard deviaties benodigde tijd bij significant verschil op interesse

Tabel

Gemiddelden en Standaard Deviaties Benodigde Tijd van Items met Significant Verschil op Interesse

	Cognitieve belasting					<i>p</i>
	<i>n</i>	Origineel		Aangepast		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
lego	67	63.72	30.56	62.16	37.42	.74
tellen	67	124.19	50.05	123.74	50.89	.96
dobbelsteen	67	37.38	17.74	41.87	25.63	.27
reeksen	67	45.94	21.41	55.19	28.30	.03
meten	67	47.57	34.00	46.33	27.13	.80